Youth Bulges and Civil Conflict: Causal Evidence from Sub-Saharan Africa

Flückiger, M., & Ludwig, M. (2018). Youth Bulges and Civil Conflict: Causal Evidence from Sub-Saharan Africa. *The Journal of Conflict Resolution* , *62*(9), 1932–1962. https://doi.org/10.1177/0022002717707303

Published in:
The Journal of Conflict Resolution

Document Version:
Peer reviewed version

Youth Bulges and Civil Conflict: Causal Evidence from Sub-Saharan Africa*

Matthias Flückiger† Markus Ludwig‡

May 8, 2017

Abstract

The presence of an exceptionally large youth population, i.e., a youth bulge, is often associated with an elevated risk of civil conflict. In this paper, we develop an instrumental variable approach in which the size of the youth cohorts in Sub-Saharan Africa are identified using variation in birth year drought incidence. Our results show that an increase in the size of the population group aged 15–19 raises the risk of low-intensity conflict. A one percent increase in the size of this age group augments the likelihood of civil conflict incidence (onset) by 2.3 (1.2) percentage points. On the other hand, we do not find any association between the size of the two adjacent youth cohorts, i.e. the population groups aged 10–14 and 20–24.

JEL classification: D7, J1, Q1

Keywords: Civil Conflict, Youth Bulge, Drought, Instrumental Variable Regression

*We thank Frank Krysiak, Michael Wagner, Hermione Miller-Moser, and Daniel Müller for helpful discussions and comments.

†Queen's University Belfast and University of Basel. Email: matthias.flueckiger@qub.ac.uk

‡University of Bayreuth. Email: markus.ludwig@uni-bayreuth.de

1 Introduction

Civil conflicts and the accompanying violence not only inflict severe human suffering, but also represent a main impediment to economic development. While the consequences of civil wars are undoubtedly devastating, the discussion regarding their causes is controversial. Highly disputed, for example, is the existence of a link between large youth cohorts and the risk of civil conflict, i.e., the existence of a youth bulge effect. While prominent theoretical arguments of Goldstone (1991) and Goldstone (2010) as well as empirical studies of Urdal (2006) and Collier et al. (2009) suggest the existence of a positive association between youth bulges and the likelihood of conflict, the youth bulge hypothesis is contested by Fearon (2011) and Sommers (2011). These contradicting findings reflect the challenges associated with the identification of youth bulge effects. Reverse causality, for instance, constitutes a major issue. While civil conflicts could be caused by an unusually large youth cohort, they also reduce the size of the cohort by increasing mortality and displacement (Murray et al., 2002; UNHCR, 2010). In this paper, we develop an instrumental variable approach for Sub-Saharan Africa (SSA) to overcome these issues.

For the subsequent analysis, we define the term 'youth' to encompass the ages 10–24. The available population data are grouped into 5-year age bins which implies that three age groups—namely, the population aged 10-14, 15–19 and 20–24—fall within this youth definition. We characterize the size of these three youth age groups by their respective absolute size. More specifically, we use time-detrended values thereof. This measure is trend-stationary and consequently captures transitory shocks in the cohort size, i.e., deviations from the country-specific trend. We identify fluctuations in the youth population sizes using variation in the drought incidence in the birth years of the respective age groups. We show that these birth-year drought shocks induce variation in the size of their respective age groups but do not possess explanatory power for adjacent groups. This allows for a clean separation of the effects of different age group on conflict within a single regression equation.

The argument underlying the drought–population relationship is that drought episodes affect agricultural output and income in the rainfed farming systems of SSA. This, in turn, influences infant health and survival. We therefore argue, and empirically document, that variation in birth-year drought incidence constitutes a valid instrument for cohort sizes. This allows us to analyze and isolate the effect of the population size of each youth age group, i.e., the age groups 10–14, 15–19, and 20–24, on the risk of civil conflict.

Our country-level drought measure—which can be interpreted as the share of the landmass in drought during the growing season—is based on the Standardized Precipitation-Evapotranspiration Index developed in Vicente-Serrano et al. (2010). This index represents a multiscalar drought measure which takes into account that drought conditions not only depend on precipitation alone, but are also affected by temperature and soil conditions. For our empirical analysis, we combine the population from the World Population Prospects and the drought data with the Uppsala/PRIO dataset on civil conflicts (Gleditsch et al., 2002). Our sample comprises the countries of SSA and spans the period 1960-2010.

We focus our analysis on Sub-Saharan Africa for several reasons. First, it is home to most of the least developed economies and is one of the regions most prone to civil war. Additionally, more than two thirds of the population is still employed in agriculture (CTA, 2012) of which 96 percent[1] is rainfed. Furthermore, SSA suffers from high child mortality, owing largely to malnutrition and poor access to medical care. Finally, evaluating the youth bulge hypothesis is of particular relevance to SSA. The very youthful and fast-growing population implies that SSA will experience a high degree of population pressure in the years to come, especially taking into account its rigid economic systems and the already strained labor markets (World Bank, 2009; Urdal, 2012).

Our regression results reveal that an increase in the population size of age group 15–19 raises the risk of low-intensity civil conflict but has no significant effect on high-intensity

[1]See World Bank (2008).

civil conflicts.[2] An increment of one percent increases (low-intensity) civil conflict incidence by 2. percentage points. On the other hand, changes in the size of the adjacent age groups, i.e. groups 10–14 and 20–25, do not affect the conflict risk. A possible explanation for this finding is that the age span 15–19 corresponds to the period of life in which the transition into the labor typically takes place in SSA (Filmer and Fox (2014, p.50, p.52) and Garcia and Fares (2008, p.22)). The presence of a large labor market entry cohort therefore implies a strong increase in labor supply. Due to the existing bottleneck effects in the labor market, unemployment and, consequently, the number of marginalized youth increases. This raises the (total) number of people for which participating in riots or joining armed groups becomes a relatively attractive option. That is, the pool of potential recruits is enlarged. The feasibility of insurgencies and other acts of civil conflicts is consequently facilitated. To support the plausibility of this interpretation, we use regional-level information on unemployment derived from harmonized surveys and show that the unemployment rate of 15–19 year olds is negatively associated with the occurrence of droughts in their birth year.

It is important to note that while our results suggest that an exceptionally large population group aged 15–19 triggers conflicts by inducing transitory fluctuations in the size of the pool of potential recruits, they do not imply that this age group accounts for the majority of the violent perpetrators. Furthermore, our results represent strictly local effects of variation in cohort sizes induced by drought conditions. This implies that our results do not directly produce policy implications. Variation in cohort sizes originating from other sources than droughts may have very different effects compared to the ones uncovered in our paper.

The interpretation of our estimates as causal effects hinges on our identification strategy being valid. One potential threat to our 2SLS-IV approach is that the exclusion restriction is violated. That is, that drought shocks not only affect the size of cohorts but additional characteristics. For rural India, for example, Shah and Steinberg (Forthcoming) show that drought shocks in birth years have permanent effects on cognitive abilities. Persons born in

[2] Low-intensity conflicts are defined as conflicts with more than 25 but less than 1,000 battle related deaths per year. High-intensity conflicts refer to civil wars with more than 1,000 battle-related deaths per year.

times of drought accumulate less human capital than people born in average years. While continent-wide, age-stratified data on cognitive abilities are not available, we use survey data to investigate whether such effects are detectable for Africa at the regional level. Our results indicate that there is no relationship between birth year drought shocks and human capital accumulation of the population aged 15–19 (as proxied by years of schooling and literacy). This suggests that the potential influence of droughts on cognitive abilities is not a source of bias for our 2SLS-IV estimation strategy. While it is impossible to make any definite claims regarding the validity of the exclusion restriction, it is important to note that, should drought events have long lasting effect on human capital accumulation, our estimates would likely be biased downward (i.e., toward zero). People who have accumulated more human capital have, for example, higher opportunity costs of fighting. The presence of this effect would dampen the positive relationship between the presence of youth bulges and conflict incidence.

To further underline the general validity of our results, we conduct various robustness checks. For example, we show that our results are not dependent on a specific definition of our drought measure or the use of a particular conflict database. Furthermore, we obtain qualitatively equivalent results if we use a relative (the share of youth in the adult total population) rather than our preferred absolute measure of the youth bulge. Finally, we are able to reproduce our country-level results at the regional level.

Our paper contributes to the extensive empirical literature on the determinants of civil war.[3] Closely related to our analysis are the studies that investigate the association between country-specific total population size and civil conflict (Goldstone, 1991; Collier and Hoeffler, 2002, 2004; Brückner, 2010). In contrast to these studies, we stress the importance of distinguishing between different age groups when analyzing population size.[4] Also closely linked to our study is the strand of literature that analyzes the effect of opportunity costs

[3] See Blattman and Miguel (2010) for a comprehensive review.

[4] For example, the setup of Brückner (2010) does not allow for the identification and isolation of effects of particular population groups, since the first-stage instrument affects the size of all population groups.

on the risk of civil conflicts (see, e.g., Miguel et al. (2004), Dube and Vargas (2013) or Hodler and Raschky (2014)). The findings show that reduced opportunity costs raise the risk that conflicts occur. Similarly, the qualitative work of Sommers (2003, 2011) identifies the marginalization of youth as an important issue in SSA. He, however, contests the existence of a youth bulge effect.

Further related to our paper is the microdata-based literature that analyzes the role of children and youth in armed conflicts in SSA. These studies document that the majority of the youth join armed forces voluntarily (ILO, 2003; Singer, 2006). An important motivation for joining is the lack of employment opportunities and, more broadly, the hope to escape poverty (Brett and Specht, 2004; ILO, 2003; Humphreys and Weinstein, 2008). In the context of coerced recruitment, Beber and Blattman (2013) show that there is a clear preference of rebels to recruit young adolescent soldiers, rather than children or adults. Adolescents are susceptible to indoctrination and manipulation, while at the same time being very able fighters. Finally, our analysis is also related to the studies that stress the relationship between climate conditions and civil conflict (Hsiang et al., 2011; Burke et al., 2009) as well as the literature that analyzes the association between agricultural income shocks and infant health and survival (e.g., Baird et al. (2011)).

The remainder of the paper is structured as follows: In Section 2, we discuss the reasoning behind the use of a birth year drought measure as instrument for the cohort size as well as potential mechanisms underlying the youth bulge effect. In Section 3, the empirical methodology is presented before the data are discussed in Section 4. Section 5 presents the results, Section 6 concludes.

2 Mechanisms

In this section, we first discuss the association between drought incidence and infant survival. In a second step, we outline potential mechanisms that lead to an increase in the likelihood of civil conflict associated with a surge in the size of the youth population.

Droughts and Cohort Size

In rainfed agricultural systems such as those in SSA, agricultural output and, consequently, income are largely determined by weather conditions in the growing period. Droughts, for example, lead to poor harvests, decrease average household income and can lead to a food crisis among large parts of the population that derive their livelihood from subsistence farming (Devereux, 2009). Infants and Children are particularly vulnerable to such shocks due to several, non-exclusive factors. First, a poor harvest can directly lead to malnutrition which reduces infant survival.[5] According to UNICEF (2012), one third of the child deaths in SSA can be attributed to malnutrition. Second, droughts also lead to a reduction in income in rural SSA. This negatively affects child health and decreases the probability of survival, particularly in the first months after birth (see, e.g., Baird et al. (2011)). Households affected by negative agricultural output shocks employ several coping strategies of which three of the most important are: Reduced food consumption, increase in distress maternal labor supply, and cutback in health service expenses (Devereux, 2009; Bhalotra, 2010). Any of these strategies reduces the likelihood of child survival. Third, droughts potentially reduce birth rates by, for example, increasing in the number of stillbirths (Hernández-Julián et al., 2014). Overall, the literature documents that the likelihood of infant and child survival in rural SSA is strongly influenced by prevailing drought conditions. Therefore, we expect that variation in drought intensity induces fluctuations in the size of birth cohorts. In the empirical analysis, we first provide evidence for the existence of this relationship before exploiting it in a 2SLS-IV setting.

[5] See for example Rice et al. (2000) for a review.

Youth Bulges and Civil Conflict

Statistics on battle-related deaths indicate that the majority of soldiers in SSA are between 15–44 years old (Murray et al., 2002). Often, however, the combatants are considerably younger. In a survey of armed groups from several countries, Beber and Blattman (2013), for example, document that over 20 percent of the recruits were aged 14 or younger. We therefore define the age range in which youth become attractive for recruitment to lie between 10 and 24. Consequently, an exceptionally strong increase in the cohort size of any cohort lying within this age span could increase the risk of conflict by simply increasing the size of the pool of potential (attractive) recruits.

The availability of a sufficiently large pool of potential recruits constitutes a prerequisite for the feasibility of insurgencies as Africa's rebels are almost exclusively composed of foot soldiers (Okumu and Ikelegbe, eds, 2010). The soldiers are usually recruited in the (rural) hinterland where influence of the government is weak (Herbst, 2000, 2004; Fearon and Laitin, 2003; Sommers, 2003). In addition to facilitating recruitment, a large idle youth population could also augment the baseline conflict risk through its effect on the outbreak of spontaneous, unorganized acts of violence, such as riots.

While the effect of mechanically augmenting the size of the pool of potential recruits is common to all youth cohorts, we argue, based on the literature, that population surges in the age group 15–19 are particularly likely to increase the risk of civil conflict. It is during this period of life that the transition into the labor market typically takes place in SSA (Filmer and Fox (2014, p.50, p.52) and Garcia and Fares (2008, p.22)). A surge in the size of this population group likely results in an exceptional increase in labor supply. The rigid labor markets in SSA, dominated by subsistence agriculture, are unlikely to be able to absorb the supply side shock (Garcia and Fares, 2008; Chandrasekhar et al., 2006). The result is a bottleneck effect characterized by high unemployment among the labor market entry cohort which, in turn, raises the number of marginalized youth. This lowers the (average) opportunity cost of participating in conflicts among the youth aged 15–19. In addition to this bottleneck effect, which is specific to the labor market entry cohort, the boost in labor supply could also deteriorate the labor market opportunities of older age groups. In that case, their opportunity costs of participating in civil conflict are also reduced. Taken together, a rise in the size of the population aged 15–19 is likely to deteriorate labor market opportunities and therefore raise the number of people for which partaking in civil conflicts becomes a veritable option.

Furthermore, Beber and Blattman (2013) show that youth are most attractive for recruitment to the rebels around the age of 15. This is due to the fact that adolescents are still susceptible to indoctrination and, at the same time, possess the physical requirements to be effective fighters. Relatedly, Farrington (1986) observes that the violent crime curve peaks at teenage age.

Summing up, we expect that surges in the size of the population aged 15–19 have the greatest impact on the risk of civil conflict among the three youth age groups. An important note relates to the age composition of the persons that participate in conflicts. This composition is ex-ante unrelated to the presence of youth bulges. Rather, it depends on the relative sizes of the age cohorts that participate in the acts of civil conflict. A further point we want to stress is that the relationship between youth bulges and conflict is likely to be different in rural and urban areas. Additionally, the extent of the youth bulge effect might also depend on the sex composition of the cohort. However, due to data limitations, we cannot analyze the existence of these types of heterogeneities.[6] Our estimates will represent the average effect of the individual youth age groups on civil conflict. In the next section, we describe our empirical strategy employed to estimate these effects.

[6]Our instrument does not allow for the separation of effects of large male and female cohorts.

3 Empirical Strategy

The main challenge in identifying the effects of large youth age groups on the risk of civil conflicts is the existence of reverse causality. While youth bulges potentially increase the incidence of civil wars, raging conflicts result in death and displacement (e.g., Murray et al. (2002) or UNHCR (2010)). Due to these effects, simple OLS estimates of the effect of youth bulges on conflict are likely to be downward biased.[7] Omitted variables as well as measurement error in the dependent variable are additional potential sources of bias.

To overcome these issues, we employ a two-stage least square instrumental variable approach in which we use drought variation in the birth years of the respective age cohorts as instrument for their size. Information regarding the size of the age groups is only available as 5-year bins. That is, the population sizes are reported for the age groups 0–4, 5–9, 10–14, 15–19, and so on. This implies that we cannot identify the effect of a specific age cohort on the risk of civil conflicts. Rather, our estimates will represent the average effect of the cohorts grouped together in 5 year bins. In this study, we focus on the analysis of the effect of the three youth age groups 10–14, 15–19 and 20–24 on the likelihood of civil conflict.

Methodology

The first stage regressions, in which we identify the size of youth cohorts using variation in the incidence of drought in their birth years, can be represented as:

$$\log P^y_{i,r,t} = \beta\, S^y_{i,r,t} + \sum_{g \neq y} \psi_g\, S^g_{i,r,t} + \alpha_{i,r} + \gamma_{i,r}\, t + T_{r,t} + v_{i,r,t}. \tag{1}$$

The dependent variable $P^y_{i,r,t}$ is the absolute population size of the youth age group y in year t and country i that is part of region r. That is, $P^y_{i,r,t}$ is either the size of age group 10–14, 15–

[7]The direction of bias when using relative measures of youth bulges, i.e., the size of the youth bulge relative to total population (employed, for example in Collier and Hoeffler (2004) or Urdal (2006)) is ambiguous. The issues related to reverse causality, however, are potentially compounded. This is due to the fact that fluctuations in any age group change the size of the relative youth bulge. This issue is discussed in more detail in Appendix A.

19, or 20–24. The birth-year specific drought measure for age group y is represented by $S^y_{i,r,t}$. Its construction will be described in detail in the next section. The drought measures for the two remaining youth cohorts, i.e., $g \neq y$, are represented by the sum given in the second term on the RHS of Eq.(1). A necessary condition for the validity of our identification strategy is that the birth-year drought measure of age group y affects its size while being uncorrelated with the population size of other age groups. This allows for a clean identification of the size of the respective youth groups.

The inclusion of country-fixed effects $\alpha_{i,r}$ in the setup implies that we will only rely on within-country variation in our regressions. The coefficients $\gamma_{i,r}$ and $T_{r,t}$ capture country-specific time trends and region-specific time fixed effects, respectively. The latter capture any region-specific changes over time, such as, for example, political stability.[8] The inclusion of country-specific time trends ensures that the population data become trend stationary, thereby avoiding spurious correlation.[9] The idiosyncratic error term is given by $v_{i,r,t}$. In all regressions, the standard errors are clustered at the country level.

To analyze the reduced-form relationship between the youth cohort specific drought measures and conflict risk, we use a regression setup analogous to Eq.(1); the only difference being that we replace the dependent variable with an indicator variable ($c_{i,r,t}$) that captures either incidence, onset or offset of civil conflict. Formally, the reduced form equation is given by:

$$c_{i,r,t} = \sum_{g \in Y} \lambda_g \, S^g_{i,r,t} + \alpha_{i,r} + \sigma_{i,r} \, t + T_{r,t} + e_{i,r,t}. \tag{2}$$

The coefficients λ_1, λ_2 and λ_3 capture the effects of birth-year drought shocks of age group 10–14, 15–19 and 20–24, respectively.

[8] The regions are taken from the World Development Indicators and are defined as: Eastern, Middle, Northern, Southern, and Western Africa. Note that our results remain stable when replacing the region-specific year fixed effects by simple year fixed effects.

[9] On a related note, it is not advisable to employ the growth rates of the youth age groups as dependent variables in the context of our study. This is due to the fact that positive growth rates do not necessarily reflect positive—and hence exceptional—deviations from the country-specific mean. Positive growth rates are also observable when the size of the youth population is below its country-specific mean. This is the case, for example, when the youth population is recovering from a preceding negative shock.

To quantify the effect of the three youth age groups on the risk of conflict we use a 2SLS instrumental variable approach, where we use the age-group specific birth-year drought shocks as instrument for their respective size. That is, in one regression we have three endogenous variables which are instrumented by three drought shocks. The second stage regression is characterized by:

$$c_{i,r,t} = \sum_{g \in Y} \theta_g \log P^g_{i,r,t} + \tau_{i,r} + \gamma_{i,r}\, t + T_{r,t} + u_{i,r,t}. \tag{3}$$

Again, $c_{i,r,t}$ is a binary variable, either representing incidence, onset or offset of civil conflict in country i and year t. The (log) size of the population youth age group g is represented by $\log P^g_{i,r,t}$. Due the inclusion of the country-specific time trends and fixed effects, our 2SLS-IV estimates will represent the effect of transitory shocks in youth population—i.e., effects of deviations from the country-specific time trend—on the risk of conflict.

It is important to note that the results of our 2SLS regressions represent local effects of variation in cohort sizes induced by drought conditions. The insights derived in this paper therefore do not carry over to changes in cohort sizes that are caused by alternative factors, such as government programs targeted at reducing fertility rates. This also means that we cannot deduce policy implications from our results. Variation in cohort sizes originating from other sources than droughts may have very different effects compared to the ones unveiled in our paper.

Threats to Identification

The 2SLS-IV setup outlined above allows for the identification of the individual youth age group effects under the assumption that (a) our cohort specific drought measure only influences the size of the respective cohort and does not affect the size of other cohorts, (b) the drought measure is exogenous to other factors influencing civil conflicts, and (c) the variation in the drought measure only affects civil conflict through the fluctuations induced in the cohort sizes.

In the empirical part, we will demonstrate that assumption (a) holds. Condition (b) is also plausibly fulfilled as the drought index only depends on variables that are exogenous to human activity. The plausibility of condition (c), on the other hand, is less clear. For it to hold, we need that historical drought shocks do not affect contemporaneous socioeconomic variables other than through their effect on cohort sizes. Specifically, we require that cohort size unrelated effects of birth year drought shocks have dissipated by the time a cohort reaches age 10, i.e., the time at which it starts influencing the size of the age group 10–14. Below, we provide supporting evidence for the plausibility of this notion.

As discussed in the introduction, a further concern is that the variation in our birth-year drought measure not only affects population size, but also influences socioeconomic characteristics, such as cognitive abilities, of these cohorts. For example, absence of droughts might increase cognitive abilities and lead to improved labor market opportunities (e.g., Shah and Steinberg (Forthcoming); Maccini and Yang (2009); Alderman et al. (2006)). The presence of this effect would likely dampen the positive relationship between the presence of youth bulges and conflict incidence. Unfortunately, we are not able to disentangle these effects due to the lack of data. It is therefore important to note that our results represent the average effect of youth bulges—including the effect of the size itself as well the potential quality effects induced by changes in cohort sizes described above—on the risk of conflict. However, in Section 5.1 we provide suggestive evidence that human capital related effects are unlikely to influence our estimates.

Before presenting the regression results we next describe the data.

4 Data

Our sample consists of country-level observations for 43 SSA nations covering the period 1960 to 2010.[10] Information regarding civil conflicts is extracted from the PRIO/Upsalla database. We only consider conflicts between a government and one or more internal opposition groups, i.e., conflicts that are fought in a country's own territory.[11] We distinguish between three types of conflicts: Any civil conflict, low-intensity civil conflict (at least 25 but less than 1,000 annual battle deaths), and high-intensity civil conflict (more than 1,000 battle deaths in a year). For each of these conflict types we construct an incidence indicator. This variable takes the value one in years of conflict, and zero otherwise.

As additional dependent variables we construct an onset and an offset indicator for the respective conflict types. The onset variable takes the value one in the first year of conflict, and zero otherwise. Following Collier and Hoeffler (2004), we code all observations of continued conflict as missing values. The offset indicator is set to one in the first year of peace, and zero otherwise. In analogy to the onset indicator, we drop years of continued peace from our analysis (see e.g., Bazzi and Blattman (2014) or Beck and Katz (2011)). In Appendix S1.11, Table S1.11, we show that our results are robust with respect to the use of alternative onset and offset definitions.

The data on cohort-specific population sizes are drawn from the World Population Prospects (2010 Revision). As mentioned in Section 3, this information is reported for individual 5-year bins. Specifically, the population sizes are reported for the age groups 0–4, 5–9, 10–14, 15–19,...., 95–99, and 100+. In our analysis, we will make use of the information regarding the population sizes of age groups 10–14, 15–19, and 20–24. Information with respect to live birth rate and infant mortality rate are derived from the World Development Indicators.

To investigate the mechanisms underlying the youth bulge effect in more detail, we use

[10] Our sample includes countries of mainland Africa plus Madagascar. For small island states, drought shock and conflict data are unavailable.

[11] This corresponds to the type 3 and 4 conflicts in the PRIO/Upsalla database. The variables sideA and location coincide for these two conflict types.

information contained in the IPUMS-international database (Minnesota Population Center, 2004). This data source provides harmonized, regionally representative surveys for a number of SSA countries. From these surveys, we construct cohort-specific variables on population size, unemployment rates and school attainment at a sub-national level.[12] The resulting dataset is an unbalanced panel that spans the years 1989–2010 and contains 136 sub-national regions that lie within 11 countries.

Country-Specific Drought Measure

The basis for our age-group specific drought measure is the Standardized Precipitation-Evapotranspiration Index (SPEI) developed in Vicente-Serrano et al. (2010). This index represents a multiscalar drought measure that takes into account that drought conditions do not only depend on precipitation alone, but also on temperature and soil conditions. It outperforms other commonly used drought measures in its ability to capture the impact of drought conditions on agricultural yield (Vicente-Serrano et al., 2012). The SPEI is available at a spatial resolution of 0.5×0.5 degrees and is reported in monthly intervals for the period 1901–2013.[13] The SPEI is reported for different time-scales which represent the number of months for which the index accounts for accumulated water deficits. Short time scales (around 3–4 months) capture seasonal conditions which most directly influence agricultural output. In our main analysis, we use the SPEI at a 4 months scale.[14] For a detailed discussion of the motivation underlying the definition of our drought intensity measure as well as the stability of our results with respect to modifications in any of the steps involved in the construction process outlined below, please see Appendix B.

[12] These sub-national regions are defined by IPUMS. They do not necessarily correspond to common levels of sub-national aggregation, such as, for example, administrative regions. For more details regarding the construction of this dataset see Appendix S2.

[13] The data are available at `http://sac.csic.es/spei/database.html`.

[14] This time scale is also employed by Harari and Ferrara (2015).

The drought measure at the country-year level ($d_{i,t}$) is defined as:

$$d_{i,t} = \frac{\sum\limits_{c \in i} \frac{\sum\limits_{m \in GP_c} s_{c,t,m}}{NGP_c}}{NC_i}. \tag{4}$$

$s_{c,t,m}$ represents an indicator variable which takes the value one if a given 0.5×0.5 degree grid cell c experiences a drought shock in month m of year t. A drought shock is defined as a SPEI value that lies below the 25th percentile of the long-term grid-cell, month-specific SPEI realizations.[15] This definition of a negative moisture shock is similar to the ones used in Burke et al. (2015), Shah and Steinberg (Forthcoming), or Harari and Ferrara (2015). To obtain a drought measure at the country level from the grid-cell specific monthly shock indicators, we aggregate over two dimensions. First, we determine the number of annual drought shocks for each grid cell. To take into account that agricultural production is most vulnerable to drought shocks during the growing period we only sum over the months that belong to the growing season (see, e.g., Harari and Ferrara (2015)). That is, we include realizations of the shock indicator if month m falls into the growing period (GP_c). A month is defined as belonging to the growing period whenever average temperature exceeds 5°C and precipitation plus moisture stored in the soil exceed half the potential evapotranspiration (FAO, 1978).[16] The number of drought shocks is then divided by the cell-specific number of months that comprise the growing period (NGP_c). The result is an index that reflects the share of total growing period months during which a cell experiences a drought shock in a given year. The second dimension over which we aggregate is space. For each country and year, we compute the mean of the yearly cell-level measures over the grid cells that lie

[15] The indicator variable $s_{c,t,m}$ takes the value one whenever the SPEI value lies below the 25th percentile of the grid-cell, month-specific SPEI realizations of the time period 1935–2010. This period reflects the time span between the earliest birth year of the age group 20–24 in the year 1960 (the start year of our sample) and the last year covered in our dataset.

[16] We obtain very similar results if we sum over all the shocks, i.e., if we also include the shocks that do not fall into the growing period (Table S1.8), Appendix S1). Compared to the results in the main part, these coefficients are estimated with less precision. This indicates including less relevant moisture shocks introduces noise in our instruments.

within the home-country territory. That is, we sum the drought shocks over all cells that lie within country i and year t and divide this sum by the number of grid cells over which country i spans (NC_i). The measure that results from these aggregation procedures can be interpreted as the share of grid cells—i.e., total landmass—in drought during their growing period in a given country and year.

On the basis of the country-year drought measure, we define our age-group specific birth year drought measure. As mentioned above, the size of these age groups are reported as 5-year bins. A straightforward way to define an age-group specific measure would be to take the average over the drought measures in the five birth years of the respective age group. However, because moisture conditions in the growing season materialize with a time lag, we expect, and empirically document, that the year prior to birth is also important for child health and survival. To avoid unnecessary serial correlation, we therefore do not include information regarding drought shocks of the last (youngest) birth year belonging to a given age group in the construction of our age-group specific instrument. Illustrated for the age group 15–19, our country-level age-group specific drought measure is:

$$S_{i,t}^{15-19} = \frac{1}{4}\left(d_{i,t-16} + d_{i,t-17} + d_{i,t-18} + d_{i,t-19}\right). \tag{5}$$

The drought shock of the youngest cohort ($d_{i,t-15}$), i.e. drought conditions in the birth year of the population aged 15, is not included in the measure. We want to stress that our results are not driven by the omission of the latest birth-year shock. It merely allows for a more precise separation of the effects of the different cohorts. In Table S1.8 we show that we obtain very similar results when we include all birth year drought shocks of an age group in the construction of the instrument. Furthermore, we also demonstrate that we are able to replicate our results using an alternative instrument that is solely based on variation in precipitation levels.

Summary Statistics

Table I depicts the summary statistics of our key variables. The average proportion of countries experiencing a conflict is approximately 18 percent, the proportion subject to an onset of conflict lies at 4 percent. Given a country is in a period of conflict, there is a 23 percent chance that it will end.

Table I: Summary Statistics Key Variables

	Mean	**Std. Dev.**	**Min.**	**Max.**	**Obs.**
	Conflict				
Incidence	0.179	0.383	0	1	2193
Onset	0.042	0.200	0	1	1890
Offset	0.226	0.419	0	1	319
	Population				
Log Population 10–14	6.401	1.349	2.433	9.849	2193
Log Population 15–19	6.247	1.346	2.235	9.704	2193
Log Population 20–24	6.078	1.341	2.039	9.584	2193
	Age-Group Specific Drought Measure				
S_t^{10-14}	0.220	0.126	0	0.779	2193
S_t^{15-19}	0.218	0.126	0	0.779	2193
S_t^{20-24}	0.208	0.120	0	0.879	2193

Notes: S_t^{10-14} represents the drought shock measure for population group aged 10–14 defined according to Eq.(5). S_t^{15-19} and S_t^{20-24} are the shocks for age groups 15–19 and 20–24, respectively.

The average size of the population group aged 10–14 is approximately 1.4 million. The average size of the two remaining youth age groups is very similar.

5 Empirical Results

In this section, we present our regression results. We first provide evidence that birth year drought shocks influence infant mortality and birth rates. Subsequently, we will demonstrate that these birth year events continue to be reflected in the population sizes of the youth age groups. We will then exploit these findings to quantify the effect of youth bulges on the risk of civil conflict in a 2SLS-IV setup.

Birth Year Drought Measure, Infant Mortality and Birth Rates

Based on the discussion presented in Section 2, we expect that drought events in the birth years increase infant mortality and decrease birth rates. To test for the existence of these effects, we regress contemporaneous mortality and birth rates on our country-year specific drought measure $d_{i,t}$ (given in Eq.(4)).

Table II: Birth Year Drought Measure, Infant Mortality and Birth Rate

	Infant Mortality Rate$_t$	Birth Rate$_t$
	(1)	(2)
d_t	2.877**	-0.569*
	(1.191)	(0.326)
d_{t-1}	2.222*	-0.342
	(1.198)	(0.281)
d_{t-2}	1.385	-0.161
	(1.473)	(0.268)
d_{t-3}	1.636	-0.001
	(1.411)	(0.248)
d_{t-4}	1.723	0.102
	(1.379)	(0.273)
Obs.	1,982	1,982
RMSE	6.694	1.378

Notes: FE estimator regressions in all columns with country dummies, country-specific time trends and regional time dummies. RMSE is the root mean square error. Infant mortality rate is the number of infant deaths per 1,000 live births. Birth rate is the number of live births inhabitant (1,000). d_t represents the country and birth-year specific drought measure in year t defined according to Eq.(4).

The results presented in column (1) of Table II confirm that drought shocks increase infant mortality.[17] The coefficient of the contemporaneous drought measure is statistically significant. As expected, infant mortality is also affected by the drought conditions in the year preceding birth. Column (2) shows that a contemporaneous drought shock reduces the birth rate. These two results imply that drought conditions in birth years affect cohort sizes. In the next section, we will show that these effects are still detectable years later in the size of the youth cohorts.

[17] Due to the fact that information regarding infant mortality and birth rate is not available for all countries/years, the number of observations is reduced compared to our main sample.

Birth Year Drought Measure, Youth Bulges and Conflict

Our identification strategy relies on the assumption that the age-group specific drought measures influence the size of the corresponding cohort while, at the same time, being uncorrelated with the population size of other cohorts. Columns (1)–(4) in Table III document that our drought measure fulfills this criteria. An increase in the age-group specific drought measure reduces the size of its corresponding age group. This effect is statistically significant at conventional confidence levels for all three youth age groups. The magnitude of the drought effect is substantial. For example, the coefficient of -0.133 (first row column (1)) implies that the population of age group 10–14 decreases by 0.13 percent when our drought measure increases by one percent.[18]

Table III: Birth Year Drought Measure, Youth Bulges and Conflict

	Log Population 10–14$_t$	Log Population 15–19$_t$	Log Population 20–24$_t$	Log Remaining Population$_t$	Conflict Incidence$_t$	
	(1)	(2)	(3)	(4)	(5)	(6)
S_t^{10-14}	-0.133***	-0.069	-0.026	0.007	-0.179	
	(0.046)	(0.045)	(0.042)	(0.050)	(0.139)	
S_t^{15-19}	-0.086	-0.183***	-0.108	-0.029	-0.414***	-0.418***
	(0.064)	(0.059)	(0.065)	(0.088)	(0.131)	(0.126)
S_t^{20-24}	-0.016	-0.039	-0.155***	-0.021	-0.080	
	(0.089)	(0.056)	(0.055)	(0.089)	(0.132)	
Obs.	2193	2193	2193	2193	2193	2193
RMSE	0.063	0.059	0.063	0.059	0.278	0.278

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. S_t^{10-14} represents the drought shock measure for population group aged 10–14 defined according to Eq.(5). S_t^{15-19} and S_t^{20-24} are the shocks for age groups 15–19 and 20–24, respectively. FE estimator regressions in all columns with country dummies, country-specific time trends, regional time dummies and robust standard errors clustered at the country level in parentheses. RMSE is the root mean square error.

Crucially, the results in Table III also document that a given age-group specific drought shock is not statistically significantly related to the size of other youth cohorts. Furthermore, the point estimates presented in column (4) document that the drought measures for the three youth age groups are unrelated to the size of the remaining population.[19] Together, these results show that we are able to identify (and separate) the size of each youth age group

[18] $\left(e^{-0.133/100} - 1\right) \times 100$

[19] The remaining population is computed as:
Total population$_t$–(Population 10–14$_t$ + Population 15–19$_t$ + Population 20–24$_t$).

using variation in their age-group specific birth year drought intensity. Before exploiting this fact in a 2SLS-IV regression, we first analyze the reduced-form effects of our drought measure on the incidence of civil conflict.

The estimates in column (5) reveal that only the drought measure of the population aged 15–19 influences the risk of civil conflict. The coefficient of this age group is significant at the 99 percent confidence level and statistically significantly larger (in absolute terms) than the coefficients of the effects of the two adjacent youth age groups.[20] As shown in Figure S1.1 of Appendix S1, we obtain the same result when running the reduced-form regression with the individual birth year drought shocks, rather than the measure grouped into 5-year bins, as explanatory variables. Only birth year drought shocks with a lag of 15 to 19 years statistically significantly influence the risk of conflict. Furthermore, we show that there is no association between our age-group specific drought measure and civil conflict incidence for any of the age groups between 25–50 (Figure S1.2).

The point estimate for the birth-year specific drought shocks of age group 15–19 remains stable when dropping the drought measures for the two adjacent age groups from the regression (column (6)). This indicates that its significance is not driven by any cross-correlation between the different age-group specific drought measures. Together with the findings in columns (1)–(4), this result implies that only an increase in the size of the population group aged 15–19 raises the risk of civil conflict. As discussed in Section 2, a likely mechanism underlying this finding is that the transition into the labor market predominantly takes place during this period of life. In Section 5.1, we provide suggestive evidence for the plausibility of this notion.

A natural concern with our approach is that the age-group specific drought measure is correlated with contemporaneous socioeconomic outcomes that are not related to population cohort sizes but influence the likelihood of civil conflict through other channels. If such

[20]The p-values for the t-test that S_t^{15-19} coefficient is smaller than the S_t^{10-14} coefficient (S_t^{20-24} coefficient) is 0.067 (0.035).

population size unrelated effects were correlated with our drought measure, the coefficient estimates of the youth bulge effects would be biased. As shown in Table S1.1 in the appendix, bias of this type seems unlikely to invalidate our results as we do not find any significant association between our drought measure for the youth age groups and various contemporaneous covariates that are commonly associated with civil conflicts. This suggests that drought shocks disperse over time and do not explain variation in aggregate, macroeconomic, outcome variables after 10 or more years. In Section 5.1 we turn to a more localized analysis and investigate whether outcome variables of specific age groups are influenced by drought shocks in their corresponding birth years.

A final concern is that our linear estimation method fails to capture potentially existing nonlinear relationships. As indicated by the nonparametric local polynomial estimates in Figure I, this is not the case. Both, the first-stage association between S_t^{15-19} and the size of the population aged 15–19 (Panel (A)) as well as the reduced-form relationship between S_t^{15-19} and conflict incidence (Panel (B)) are monotonically decreasing and approximately linear. In the next step, we will quantify the youth bulge effects using a 2SLS-IV approach.

(A) (B)

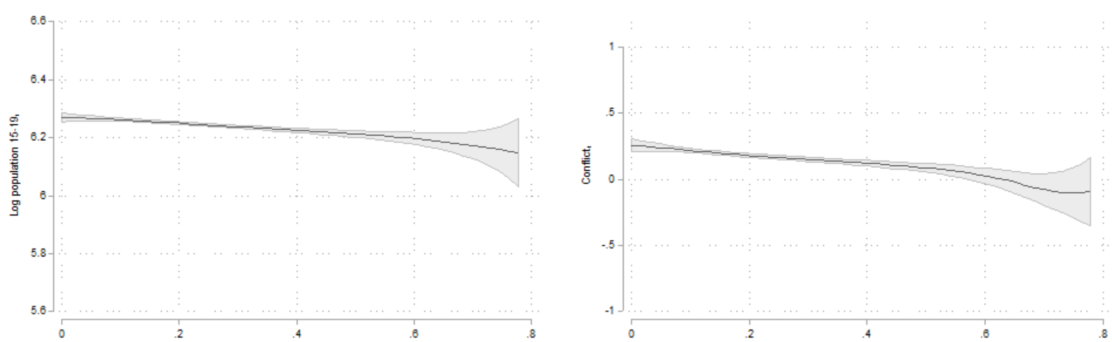

Figure I: Panel(A): Drought shock measure age group 15–19 and log population 15–19. Panel (B): Drought shock measure age group 15–19 and civil conflict incidence. Drought shock measures are defined according to Eq.(5). Estimates are produced employing the semiparametric regression approach of Robinson (1988) with a second degree Epanechnikov kernel. The parametric part of the regression model includes country and regional time fixed effects as well as country-specific time trends. The shaded areas represent the 95 percent confidence bands.

Instrumental Variable Approach

Table IV presents the 2SLS-IV estimates of the effect of changes in the size of the three youth age groups—instrumented with their age-group specific drought measure—on the risk of civil conflict. For the endogenous variables, we report the p-values of the Anderson-Rubin Chi-squared test statistic (in brackets).[21] This test statistic is robust to weak instruments (Andrews and Stock, 2005). The first-stage F-statistics of the excluded instruments as well as the p-values of the Kleibergen-Paap Chi-squared underidentification test statistic indicate that all our models are well identified. Following the practice in the literature, we will control for contemporaneous drought shocks as well as the first and second lag thereof in all regressions.[22]

Table IV: Youth Bulges and Civil Conflict Incidence: 2SLS-IV

	Civil Conflict Incidence$_t$			High Intensity Conflict$_t$	Low Intensity Conflict$_t$	Low Intensity Incidence$_t$	Low Intensity Onset$_t$	Low Intensity Offset$_t$
	(OLS)	(IV)		(IV)	(IV)	(IV)	(IV)	(IV)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log Population 10–14$_t$	0.090 (0.259)	0.280 [0.787]						
Log Population 15–19$_t$	0.071 (0.225)	2.188** [0.042]	2.284*** [0.000]	-0.049 [0.909]	2.343*** [0.000]	1.804*** [0.000]	1.178*** [0.000]	0.144 [0.964]
Log Population 20–24$_t$	-0.250 (0.251)	-0.062 [0.931]						
Incidence Civil Conflict$_{t-1}$						0.297*** (0.044)		
Contemporary Drought Shocks	yes	yes	yes	yes	yes	yes	yes	yes
Obs.	2193	2193	2193	2193	2193	2193	1890	319
RMSE	0.280	0.315	0.313	0.198	0.298	0.274	0.206	0.398
F-test excl. IV		3.03, 3.51, 2.89	8.520	8.520	8.520	8.362	11.16	3.162
Kleibergen-Paap		0.038	0.023	0.023	0.023	0.024	0.013	0.030

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. FE estimator regressions in all columns with country dummies, country-specific time trends, regional time dummies and robust standard errors clustered at the level in parentheses. The values in brackets represent the Anderson-Rubin Chi-squared test statistic which is robust to weak instruments. In the case of multiple endogenous variables, we report the subset Anderson-Rubin test statistic for each structural parameter (Guggenberger et al., 2012). RMSE is the root mean square error. The F-test is the first-stage Angrist-Pischke test statistic of the excluded instruments. Kleibergen-Paap is the p-value of the first-stage LM statistic of the excluded instruments. The three F-test values in column (2) represent the first-stage Angrist-Pischke test statistic of the excluded instruments of the three first stage regressions. I.e, with the (log) population of the age group 10–14, 15–19 and 20–24, respectively, as the LHS variable.

[21] In the case of multiple endogenous variables, we report the subset Anderson-Rubin test statistic for each structural parameter. See, for example, Guggenberger et al. (2012) or Kleibergen (2004) for a more detailed exposition.

[22] Our results remain stable irrespective of the inclusion of these control variables. Similarly, our results remain unchanged when we control for GDP, the Polity2 democracy score, real openness, or the rate of urbanization. The results are available upon request. Due to the possibility of including 'bad controls' (Angrist and Pischke, 2009), we do not include a large set of control variables.

We start by presenting naive OLS regressions of conflict incidence on the (log) population size of youth age groups in column (1). Neither of the coefficients is statistically significant. However, when we employ the 2SLS-IV methodology in column (2), we see—as implied by the results of Table III—that the population size of age group 15–19 affects the risk of civil conflict, while the adjacent age groups are unrelated to conflict incidence. The point estimate implies that a one percent increase in the size of the population aged 15–19 raises the likelihood of civil conflict incidence by about 2.2 percentage points.[23] It is important to note that even though only an increase in population group aged 15–19 raises the risk of conflict, this does not imply that violence is nonexistent in older age groups. The result merely suggest that this age group facilitates recruitment for acts of civil conflict by increasing the pool of potential recruits.

The fact the OLS regressions produce null results while the 2SLS-IV estimates unveil the existence of a youth bulge effects suggests that reverse causality causes a downward bias in the OLS results. As mentioned above, this is due to the fact that civil conflicts increase the death rate and induce displacement. The negative sign of the coefficient of the age group 20–24 (column (1)) indicates that these effects are potentially very strong. This is in line with the results of Murray et al. (2002),who show that the battle related death toll is highest among this population group. These results document the need for an instrumental variable in order to properly identify the youth bulge effects.

In column (3) we drop the non-significant population groups from the regression setup. Compared to column (2), the estimate for the effect of age group 15–19 remains virtually unaltered. We will subsequently restrict our attention to the analysis of effects of this age group.

To investigate heterogeneities in the type of conflict that are affected by the presence of a youth bulge, we separately analyze its effect on high-and low-intensity conflicts. The estimates presented in column (4) show that the presence of an exceptionally large population

[23] $\ln\left({(100+2.188)}/{100}\right) \times 100 = 2.164$

group aged 15–19 does not statistically significantly raise the risk of high-intensity conflict. In column (5) we estimate the youth bulge effect on low-intensity conflict incidence. The point estimate implies an increase in the risk of conflict of 2.3 percentage points as a result of a one percent increase in the youth population aged 15-19.[24] Consequently, the youth bulge primarily influences the (overall) conflict incidence through its effect on low-intensity conflicts. We will subsequently focus on the analysis of this type of conflict.

In the last step of our main empirical analysis, we turn our attention to the dynamics of the youth bulge effect. First, we control for lagged conflict in order to account for the persistence of conflicts (column (6)).[25] The Population 15–19$_t$ coefficient remains statistically significant; the long-run effect of 2.566 is very similar to the point estimate presented in column (3).[26]

So far, we have used the incidence of civil conflicts as dependent variable. The variation in this indicator stems from both the onset and the offset of civil conflicts. To investigate whether youth bulges differentially affect these outcomes, we employ them separately as dependent variables in columns (7) and (8).[27] The findings with respect to onset and offset are robust to alternative variations in coding (see Table S1.11). The estimate in column (7) shows that a one percent increase in the size of the population aged 15–19 increases the risk of low-intensity civil conflict onset by 1.17 percentage points. This finding is consistent with our expectations. The presence of a youth bulges increases the pool of potential recruits. Therefore, recruitment is easier for rebel armies which allows them to gain enough strength to instigate civil conflict. However, we do not detect any effect on the duration, i.e., the offset of conflicts. A possible interpretation of this result is that insurgencies which arise due to the increase in the size of potential recruits could be (relatively) unorganized and based

[24]Evaluated at the mean, a one percent increase in the population aged 15–19 raises the low-intensity conflict risk by 12 percent.

[25]The fact that or sample covers 50 years significantly reduces the extent to which the Nickel bias could influence our regression results when we include the lagged dependent variable as a regressor (Nickell, 1981).

[26]The long-run effect is given by 1.804/(1–0.297).

[27]For the analysis of the onset (offset), we restrict the sample to the observations which are at risk of transitioning into periods of civil conflict (peace). Consequently, we exclude all observations with periods of continued conflict (peace) (see, e.g., Bazzi and Blattman (2014)).

on largely inexperienced fighters.

Before turning to the final step of our analysis, we want to note that the results presented above are stable to a variety of robustness checks (see Appendix S1). Particularly, we document that the estimates are robust to modifications in our drought measure as well as to the use of a rainfall instrument. Furthermore, we show that we obtain very similar results when employing individual birth year drought shocks (rather than the 5 year age-group drought measure) as instruments for the cohort sizes. Additionally, we replicate our results using coup attempts rather than conflict incidence as dependent variable. Finally, we demonstrate that we obtain qualitatively equivalent results if we use a relative measure of youth bulge (youth as share in total adult population) instead of the absolute measure (Table S1.12).

5.1 Mechanisms

As outlined in the discussion of potential mechanisms in Section 2, there are two potentially important channels through which the existence of a youth bulge influences the risk of civil conflict. First, a positive shock in the size of population group aged 15–19 directly implies an enlargement of the population pool from which rebels typically recruit. Second, there is potentially a strong increase in labor supply as the youth often transit into the labour market between the ages 15–19 (Filmer and Fox (2014, p.50, p.52) and Garcia and Fares (2008, p.22)). If the result is increased risk of unemployment, this would reduce the opportunity costs of partaking in civil conflict. To investigate the existence of a labor market effects, we construct a regional (sub-national) level dataset on cohort-specific population sizes, employment status and educational attainment from sub-nationally representative individual-level data drawn from the IPUMS database. The survey-based data are only available for a limited number of SSA countries and years.[28] The results do not necessarily carry over to countries not included in this limited sample. Therefore, the subsequently presented results should be

[28]Please refer to Appendix S2 for more details on this dataset.

interpreted as suggestive evidence. The regression setup is analogous to the one described in Section 3, with the exception that our unit of analysis no longer is a country but a region. In all regressions, we control for contemporaneous drought shocks and the first and second lag thereof as well as country-specific time fixed effects. The inclusion of the latter implies that we only exploit within country-year variation in our regressions.

Table V: Mechanism

	Log Population 15–19$_t$	Conflict Incidence$_t$		Unemployment Rate Population 15–19$_t$		Unemployment Rate Population 15–65$_t$		Log Years of Schooling Population 15–19$_t$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
S_t^{15-19}	-0.314**	-0.535**		-0.150**		-0.003		-0.019	
	(0.129)	(0.251)		(0.071)		(0.019)		(0.109)	
Log Population 15–19$_t$			1.704*		0.479**		0.010		0.066
			[0.063]		[0.014]		[0.861]		[0.882]
Obs.	388	388	388	388	388	388	388	336	336
RMSE	0.124	0.275	0.417	0.069	0.113	0.014	0.017	0.514	0.651
F-test excl. IV			5.573		5.573		5.573		4.441
Kleibergen-Paap			0.019		0.019		0.019		0.035

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. S_t^{15-19} represents the drought shock measure for population group aged 15–19 defined according to Eq.(5). FE estimator regressions in all columns with country dummies, country-specific time trends, regional time dummies and robust standard errors clustered at the level in parentheses. The values in brackets represent the Anderson-Rubin Chi-squared test statistic which is robust to weak instruments. RMSE is the root mean square error. The F-test is the first-stage Angrist-Pischke test statistic of the excluded instruments. Kleibergen-Paap is the p-value of the first-stage LM statistic of the excluded instruments.

The first three columns in Table V document that drought shocks in the birth years of the population group aged 15–19 influence its size as well as the risk of conflict. These results confirm our findings of the country-level analysis. In Appendix S2 we show that drought shocks in the birth years of adjacent cohorts have no effect on any of the dependent variables used in Table V. In column (3) we instrument population size with birth year drought shocks. The coefficient of 1.7 implies that a one percent increase in the population size of the age group 15–19 raises the likelihood of conflict by about 1.7 percentage points. This estimate is similar, albeit somewhat lower, compared to the coefficient resulting from country-level regressions (Table IV, column (3)).

The results presented in columns (4) and (5) provide suggestive evidence for the importance of labor market effects in explaining the positive association between youth bulges and conflict risk. Droughts in birth years reduce cohort sizes which, in turn, reduce the number of people that transit into the labor market 15 to 19 years later. This explains the negative reduced-form effect (column (4)). In column (5), we quantify this effect using the 2SLS-IV approach. The resulting estimate suggests that a one percent increase in the population aged 15–19 increases unemployment among this group by about 0.5 percentage points. On the other hand, we do not find that the presence of a youth bulge increase the overall unemployment rate among the working age population (columns (6)–(7)).

As a final step of our analysis, we use the regional-level data to investigate whether we observe birth year drought shock related differences in human capital accumulation, as proxied by years of schooling. Our regression results suggest that this is not the case. The point estimates of both the reduced-form and 2SLS-IV regression are statistically not significant at conventional significance levels and close to zero.[29] This null-result mitigates concerns related to the possibility that birth year drought shocks influence human capital accumulation and thereby invalidates the exclusion restriction central to our identification strategy.[30]

Taken together, our empirical analysis demonstrates that an increase in the size of the population aged 15–19 increases the risk of civil conflict. A likely mechanism underlying this finding is a deterioration in labor market opportunities for youth cohorts which leads to a decrease in the opportunity costs of partaking in conflicts.

6 Conclusion

Youth bulges are often named as an important determinant of civil conflict, both in theoretical and empirical work. So far, however, causal evidence for the existence a youth bulge effect is missing due to issues of reverse causality and omitted variables. To address these shortcomings, we employ drought variation in the birth years of the respective age groups

[29]We find qualitatively equivalent results if we employ literacy rates instead of years of schooling as dependent variable.

[30]As discussed earlier, the existence of a negative drought effect on cognitive abilities or education would likely bias our results towards zero. The absence of drought shocks would increase human capital and raise the opportunity cost of partaking in civil conflicts, thereby dampening the negative effects of youth bulges found above.

as an instrument for their size. Drought shocks in birth years raise infant mortality and decrease birth rates. The fact that these effects persist allows us to use age-group specific birth year drought shocks as instruments for their population size.

Our estimates reveal a substantial effect of transitory shocks in the size of the population group aged 15–19 on the risk of civil conflict in Sub-Saharan Africa. An increase in population of one percent raises the risk of low-intensity civil conflict incidence by 2.3 percentage points. On the other hand, we do not find any association between conflict incidence and the size of either adjacent youth age group, i.e., age group 10–14 and 20–24. A possible interpretation for the crucial importance of population group aged 15–19 is that the transition into the labor market takes place during this period of life. In the presence of an exceptionally large labor market entry cohort, the labor markets are likely to experience a bottleneck effect and, consequently, lead to an increase in the number of marginalized youth. Their opportunity cost of partaking in civil conflict is consequently reduced. Our results therefore suggest that inducing structural changes which facilitate the transition into the labor market would likely mitigate the negative effects associated with the occurrence of youth bulges.

References

Alderman, Harold, John Hoddinott, and Bill Kinsey (2006) 'Long term consequences of early childhood malnutrition.' *Oxford Economic Papers* 58(3), 450–474

Andrews, Donald W.K., and James H. Stock (2005) 'Inference with weak instruments.' *NBER Technical Working Papers 0313*

Angrist, Joshua, and Jörn-Steffen Pischke (2009) *Mostly Harmless Econometrics: An Empiricist's Companion* (Princeton University Press)

Baird, Sarah, Jed Friedman, and Norbert Schady (2011) 'Aggregate income shocks and infant mortality in the developing world.' *The Review of Economics and Statistics* 93(3), 847–856

Bazzi, Samuel, and Christopher Blattman (2014) 'Economic shocks and conflict: Evidence from commodity prices.' *American Economic Journal: Macroeconomics* 6(4), 1–38

Beber, Bernd, and Christopher Blattman (2013) 'The logic of child soldiering and coercion.' *International Organization* 67, 65–104

Beck, Nathaniel, and Jonathan N. Katz (2011) 'Modeling dynamics in time-series–cross-section political economy data.' *Annual Review of Political Science* 14(1), 331–352

Bhalotra, Sonia (2010) 'Fatal fluctuations: Cyclicality in infant mortality in India.' *Journal of Development Economics* 93(1), 7–19

Blattman, Christopher, and Edward Miguel (2010) 'Civil war.' *Journal of Economic Literature* 48(1), 3–57

Brett, Rachel, and Irma Specht (2004) *Young Soldiers: Why They Choose to Fight* (Lynne Rienner Publishers)

Brückner, Markus (2010) 'Population size and civil conflict risk: Is there a causal link?' *Economic Journal* 120(544), 535–550

Burke, Marshall B., Edward Miguel, Shanker Satyanathd, John A. Dykemae, and David B. Lobell (2009) 'Warming increases the risk of civil war in africa.' *Proceedings of the National Academy of Sciences USA* 106, 20670–20674

Burke, Marshall, Erick Gong, and Kelly Jones (2015) 'Income shocks and hiv in africa.' *The Economic Journal* 125(585), 1157–1189

Chandrasekhar, C. P., Jayati Ghosh, and Anamitra Roychowdhury (2006) 'The 'demographic dividend' and young India's economic future.' *Economic and Political Weekly* 41(49), 5055–5064

Collier, Paul, and Anke Hoeffler (2002) 'On the incidence of civil war in Africa.' *Journal of Conflict Resolution* 46(1), 13–28

_ (2004) 'Greed and grievance in civil war.' *Oxford Economic Papers* 56(4), 563–595

Collier, Paul, Anke Hoeffler, and Dominic Rohner (2009) 'Beyond greed and grievance: Feasibility and civil war.' *Oxford Economic Papers* 61(1), 1–27

Coulibaly, Ibrahima (2004) 'The impact of HIV/AIDS on the labour force in Sub-Saharan Africa: A preliminary assessment.' *International Labour Organization: Research and policy analysis*

CTA (2012) 'Increasing rural employment in Sub-Saharan Africa.' *CTA Policy Brief*

Devereux, Stephen (2009) 'Seasonality and social protection in Africa.' *FAC Working Paper No. SP07*

Dube, Oeindrila, and Juan Vargas (2013) 'Commodity price shocks and civil conflict: Evidence from Colombia.' *The Review of Economic Studies*

FAO (1978) 'Report on agro-ecological zones project. Vol.1: Methodology and results for Africa.' World Soil Resources Report 48, Food and Agriculture Organization of the United Nations

Farrington, David P. (1986) 'Age and crime.' *Crime and Justice* 7, 189–250

Fearon, James D. (2011) 'Governance and civil war onset.' *World Bank, Washington, DC*

Fearon, James D., and David D. Laitin (2003) 'Ethnicity, insurgency, and civil war.' *American Political Science Review* 97(1), 75–90

Filmer, Deon, and Louise Fox (2014) *Youth Employment in Sub-Saharan Africa* Africa Development Forum (World Bank Publications)

Garcia, Marito, and Jean Fares (2008) *Youth in Africa's Labor Market* (The International Bank for Reconstruction and Development / The World Bank)

Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg, and Havard Strand (2002) 'Armed conflict 1946-2001: A new dataset.' *Journal of Peace Research* 39(5), 615–637

Goldstone, Jack A. (1991) *Revolution and Rebellion in the Early Modern World* (University of California Press)

_ (2010) 'The new population bomb.' *Foreign Affairs*

Guggenberger, Patrik, Frank Kleibergen, Sophocles Mavroeidis, and Linchun Chen (2012) 'On the asymptotic sizes of subset anderson-rubin and lagrange multiplier tests in linear instrumental variables regression.' *Econometrica* 80(6), 2649–2666

Harari, Mariaflavia, and Eliana La Ferrara (2015) 'Conflict, climate and cells: A disaggregated analysis.' Technical Report

Herbst, Jeffrey (2000) *States and Power in Africa: Comparative Lessons in Authority and Control* (Princeton University Press)

_ (2004) 'African militaries and rebellion: The political economy of threat and combat effectiveness.' *Journal of Peace Research* 41(3), 357–369

Hernández-Julián, Rey, Hani Mansour, and Christina Peters (2014) 'The effects of intrauterine malnutrition on birth and fertility outcomes: Evidence from the 1974 Bangladesh famine.' *Demography* 51(5), 1775–1796

Hodler, Roland, and Paul A. Raschky (2014) 'Economic shocks and civil conflict at the regional level.' *Economics Letters* 124(3), 530–533

Hsiang, Solomon M., Kyle C. Meng, and Mark A. Cane (2011) 'Civil conflicts are associated with the global climate.' *Nature* 476, 438–441

Humphreys, Macartan, and Jeremy M. Weinstein (2008) 'Who fights? The determinants of participation in civil war.' *American Journal of Political Science* 52(2), 436–455

ILO (2003) *Wounded Childhood: The Use of Children in Armed Conflict in Central Africa* (Geneva: International Labour Office)

Kleibergen, Frank (2004) 'Testing subsets of structural parameters in the instrumental variables.' *The Review of Economics and Statistics* 86(1), 418–423

Maccini, Sharon, and Dean Yang (2009) 'Under the weather: Health, schooling, and economic consequences of early-life rainfall.' *American Economic Review* 99(3), 1006–1026

Miguel, Edward, Shanker Satyanath, and Ernest Sergenti (2004) 'Economic shocks and civil conflict: An instrumental variables approach.' *Journal of Political Economy* 112(4), 725–753

Minnesota Population Center (2004) *Integrated Public Use Microdata Series, International: Version 6.3* University of Minnesota (Minneapolis)

Murray, C J L, G King, A D Lopez, N Tomijima, and E G Krug (2002) 'Armed conflict as a public health problem.' *British Medical Journal* 324, 346–349

Nickell, Stephen J (1981) 'Biases in dynamic models with fixed effects.' *Econometrica* 49(6), 1417–26

Okumu, Wafula, and Augustine Ikelegbe, eds (2010) *Militias, Rebels and Islamist Militants, Human Insecurity and State Crisis* (Institute for Security Studies)

Rice, Amy L., Lisa Sacco, Adnan Hyder, and Robert E. Black (2000) 'Malnutrition as an underlying cause of childhood deaths associated with infectious diseases in developing countries.' *Bulletin of the World Health Organization* 78, 1207–1221

Robinson, Peter M (1988) 'Root- N-consistent semiparametric regression.' *Econometrica* 56(4), 931–54

Shah, Manisha, and Bryce Millett Steinberg (Forthcoming) 'Drought of opportunities: Contemporaneous and long term impacts of rainfall shocks on human capital.' *Journal of Political Economy*

Singer, Peter Warren (2006) *Children at War* (University of California Press)

Sommers, Marc (2003) 'War, urbanization and Africa´s youth at risk: Understanding and addressing future challenges.' *Basic Education and Policy Support (BEPS) Activity and Creative Associates International*

_ (2011) *International Journal of Conflict and Violence*

UNHCR (2010) 'UNHCR statistical yearbook 2009.' Geneva

UNICEF (2012) 'Levels & trends in child mortality'

United Nations (2003) 'The impact of HIV/AIDS on mortality.' *Workshop on HIV/AIDS and Adult Mortality in Developing Countries. UN/POP/MORT/2003/14*

Urdal, Hendrik (2006) 'A clash of generations? Youth bulges and political violence.' *International Studies Quarterly* 50(3), 607–629

_ (2012) 'A clash of generations? Youth bulges and political violence.' *United Nations Department of Economic and Social Affairs, Population Division Expert Papers Series*

Vicente-Serrano, S. M., S. Beguería, J. I. López-Moreno, M. Angulo, and A. El Kenawy (2010) 'A new global 0.5°gridded dataset (1901–2006) of a multiscalar drought index: Comparison with current drought index datasets based on the palmer drought severity index.' *Journal of Hydrometeorology* 11(4), 1033–1043

Vicente-Serrano, Sergio M., Santiago Beguería, Jorge Lorenzo-Lacruz, Jesús Julio Camarero, Juan I. López-Moreno, Cesar Azorin-Molina, Jesús Revuelto, Enrique Morán-Tejeda, and Arturo Sanchez-Lorenzo (2012) 'Performance of drought indices for ecological, agricultural, and hydrological applications.' *Earth Interactions* 16(10), 1–27

Vrieling, Anton, Jan de Leeuw, and Mohammed Y. Said (2013) 'Length of growing period over africa: Variability and trends from 30 years of ndvi time series.' *Remote Sensing* 5(2), 982

World Bank (2008) 'World development report 2008: Agriculture for development'

_ (2009) 'Youth and unemployment in Africa: The potential, the problem, the promise.' Online

Appendices

A Youth Bulge and Conflict: Choice of Measure

In this Section, we show why a separation of the effects of surges in youth population from effects caused by fluctuations in other population groups is only achievable using an absolute, rather than a relative, measure for the size of youth bulges. For ease of exposition, we will refer to the 'youth bulges' as the total population aged 10–24. The results below, however, directly transfer to the three age subgroups (10–14, 15–19, 20–24) encompassed in this group.

Measure for Absolute Size of Youth Bulges

Assume that a country i possesses the equilibrium capacity in economic and political institutions to absorb a youth population of $\overline{X(T)}_i$, where T is the age of the respective youth population $X(T)$ under investigation. Hence, the youth population that cannot be absorbed by the institutions is given by $X(T)_{i,t} - \overline{X(T)}_i = \Psi$. A positive shock in the size of the youth population implies that $\Psi > 0$. This, in turn, implies that the number of idle youth increases which enlarges the pool of potential recruits for revolters. As a consequence, the risk of conflict rises. In terms of a linear model this can be formulated as follows:

$$c_{i,t} = \alpha\,(X(T)_{i,t} - \overline{X(T)}_i) + \beta\,x_{i,t} = \alpha\,\Psi + \beta\,x_{i,t}, \tag{A.1}$$

where $C_{i,t}$ expresses the risk of conflict and $x_{i,t}$ are additional factors that influence this risk, such as the population structure or political institutions. Now let $x_{i,t}$ move to the error term $u_{i,t}$:

$$\mathrm{c}_{i,t} = \alpha\,(X(T)_{i,t} - \overline{X(T)}_i) + u_{i,t} = \alpha\,\Psi + u_{i,t}. \tag{A.2}$$

Note that this equation corresponds to fixed effects regression model presented in Eq.(3) under the assumption that the country-specific mean represents the equilibrium number of youth. If it is possible to identify the country-specific fluctuations around the equilibrium ($E(X(T)_{i,t}u_{i,t}) = 0$), we can estimate its effect on the risk of civil conflict, holding everything else constant (including labor market conditions, population structure and political institutions). It is important to stress that this does not imply that countries with a larger population are more prone to civil conflict. Rather, it means that nations with strong positive deviations from their respective average level of youth population are more likely to be the locus of civil wars.

Measure for Relative Size of Youth Bulges

The use of a scaled youth bulge measure makes the separation between effects that are caused by fluctuations in the youth population and effects that are induced by changes in the composition of other population groups difficult. We illustrate this using the measure of youth population relative to total population. Formally, this measure is represented by $S_{i,t} = \frac{X(T)_{i,t}}{P_{i,t}}$, where $P_{i,t}$ is the total population.

Assume that the covariance between youth bulges and the risk of civil conflict is positive, i.e., $\text{Cov}(X(T)_{i,t}, C_{i,t}) > 0$. It then follows that $\text{Cov}(S_{i,t}, C_{i,t}) > 0$, since $\text{Cov}(X(T)_{i,t}, S_{i,t}) > 0$. I.e., there is a positive association between the relative youth population and the risk of civil conflict.

If we assume that $\text{Cov}(P_{i,t}, C_{i,t}|X(T)_{i,t}) \leq 0$ we have that $\text{Cov}(S_{i,t}, C_{i,t}) \leq 0$, since $\text{Cov}(P_{i,t}, S_{i,t}|X(T)_{i,t}) < 0$.[31] We therefore find a negative association between the relative size of the youth bulge and civil conflict. However, the decrease in the relative size of the youth bulge could be due to changes in other population groups, e.g., an increase in the birth rate.

Whether we find a positive or negative association between the relative size or youth popu-

[31] Such a negative correlation, for example, was caused by the HIV epidemic in SSA which led to a severe decline in the size of the population aged 30–39 (United Nations, 2003). This drop in the adult population resulted in a reduction of marginalized youth through the improvement of employment opportunities (Coulibaly, 2004). In terms of the model presented above, this implies $\text{Cov}(S_{i,t}, C_{i,t}) < 0$.

lation and conflict therefore depends on the underlying mechanism that induces the change in the relative measure. It is therefore not possible to separate the effect of youth population pressure from changes in other population groups by using a relative youth bulge measure. However, employing an absolute measure , i.e., $X(T)_{i,t} - \bar{X}_i$ as described above, allows us to hold $P_{i,t}|X(T)_{i,t}$ constant if identification of $X(T)_{i,t}$ is possible. This enables us to isolate the effect of youth bulges on civil conflict.

B Construction of Drought Measure

The basis of our age-group specific drought intensity measure is the Standardized Precipitation-Evapotranspiration Index (SPEI) developed in Vicente-Serrano et al. (2010). This index represents a multiscalar drought measure that takes into account that drought conditions do not only depend on precipitation alone, but also on temperature and soil conditions. The SPEI is available at a spatial resolution of 0.5×0.5 degrees and is reported in monthly intervals for the period 1901–2013.[32] The SPEI is reported for different time-scales which represent the number of months that are taken into account in determining the accumulated water deficits. In order to translate the continuous, grid-cell-month specific index into a country-year specific measure of drought severity, we take several steps. These are outlined below.

Nature of Independent Variable

In principle, we could define a continuous country-year specific measure of the SPEI by aggregating (e.g., averaging) over all grid-cell level SPEIs that fall into a given country and year. The problem associated with this type of measure is that within-country heterogeneities are averaged out. If one region experiences severe drought conditions while other regions see above-average precipitation, a continuous measure will give no indication of existing

[32] The data are available at `http://sac.csic.es/spei/database.html`.

droughts. On the other hand, bivariate drought shock measures, i.e., zero-one variables, are not prone to this offsetting effect. If drought shocks occur, they will contribute to the variation in country-level moisture indices.

We define a drought shock as a SPEI below the 20th percentile of the grid-cell, month-specific SPEI realization between the years 1935 and 2010 (see below for more details on this cut-off).[33]

Time scale of the SPEI

The SPEI index is available in different time scales (3–5 months scales). Short time scales (around 3–4 months) capture seasonal conditions which most directly influence agricultural output. In our main analysis, we use the SPEI at a 4 months scale.[34] In robustness checks, we show that our results remain stable when using 3 or 5 months scales instead (Tables S1.6–S1.7).

Cut-off for shock definition

We have no prior as to which percentile represents the most adequate cut-off for the definition of a drought shock. Literature that analyzes at which value the SPEI starts representing a negative shock that affects child health and survival does not exist. In order to make a (somewhat) informed decision, we analyze the effect of the SPEI on child mortality at different percentiles of the distribution. Due to the unavailability of child mortality data at the grid-cell level, we will conduct this analysis at the country level. To this end, we compute the mean value of the SPEI for each country and year. Subsequently, we identify into which decile of the distribution each SPEI realization falls (cf. Burke et al. (2015)). We then use a non-linear local polynomial regression to analyze the effect on child mortality across different percentiles of the SPEI realizations. In analogy to the linear regression

[33] This period reflects the time span between the (first) birth year of the cohort that is aged 20–24 in the year 1960 (i.e., the starting point of our sample) and the last year included in our sample. Our results remain stable to varying this time span.

[34] This time scale is also employed by Harari and Ferrara (2015).

presented in the main part, we first purge the data of country-fixed effects, country-specific time trends and regional time fixed-effects. The results presented in Figure B.1 indicate that SPEI realizations below the 20th percentile are associated with an increase in child mortality. Based on this observation, we define all SPEI realizations below the 20th decile of the grid-cell-month specific long-term distribution as a negative shock. As depicted in Tables S1.2–S1.5, we obtain similar estimates if move this cut-off by 10 percentiles in either direction.

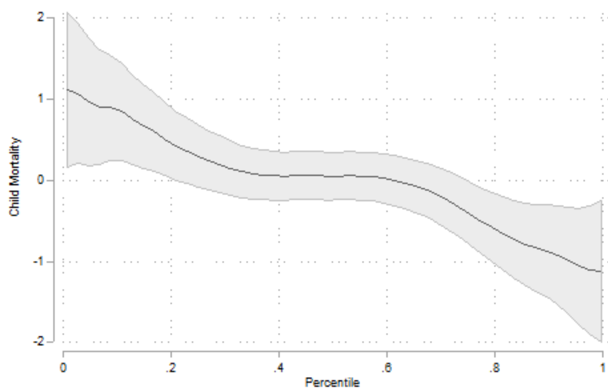

Figure B.1: Effect of SPEI on Child Mortality

Aggregation across time and space

After completing the steps outlined above, we are left with a dummy variable at the grid-cell level that indicates whether a drought occurs in a given month and year or not. In order to obtain a variable at the country-year level—our unit of analysis in the main part—we build the average over all months within a year and across all cells that lie within a given country. To gain precision we only include realizations of the SPEI into our measure that fall into the growing period. This is when crops are most vulnerable to unfavourable conditions. Including moisture conditions that fall outside the growing period into our drought measure likely introduces uninformative noise (see Table S1.8). The definition of the growing season is taken from FAO (1978). It defines a month as being part of the growing season when average temperature exceeds 5°C and precipitation plus moisture stored in the soil exceed half the potential evapotranspiration. The advantage of this measure is that it is independent of local population densities or agricultural practices. Both could change over (relatively long) the time period 1960–2010 and thereby influence our results.[35]

After averaging all the drought indices of cells that are in their growing period across time and space we obtain our final country-year level drought index. This measure represents the share of grid cells in drought during their growing period. The drought measure at the regional level, used in supporting regressions in Section 2, is computed analogously.

[35]As documented in Table S1.8, we obtain similar results when we do not restrict our measure to only include SPEI realizations that fall into the growing period (i.e. when we include all observations). While the coefficient estimates remain stable, the standard errors—as expected—increase. In Table S1.9, we additionally show that we obtain very similar estimates if we use the growing season definition of Vrieling et al. (2013) which is based satellite-derived vegetation patterns of the years 1981–2011.